A Project Proposal Titled

SMARTCLICK

_____

Submission to:

Dark Patterns Buster Hackathon
DPBH - 2023

Submitted by:

8 Bits Brawlers Team PG-BDBA
Advanced Computing Training School (ACTS),
Centre Development of Advanced Computing
(C-DAC), Bangalore

Contents

I. BACKGROUND INFORMATION

1. Title of the Project

SMARTCLICK

2. Team Leader

Ms. Sheetu Gupta

3. Team Members

Ms. Pooja Kaware, Mr. Nikhil Dubey, Mr. Ekant Yadav, Mr. Amit

4. Nodal Centre

ACTS, CDAC, Bangalore

II. TECHNICAL INFORMATION

1. Aim and Scope of the Project

The prevalence of dark patterns in e-commerce portals and apps is an issue of great concern considering huge volume of Internet users being disguised or deceived into performing actions which they don't intend to. These dark patterns impact user's trust towards Internet, its information and experience. This increasing harmful effects of the dark patterns encourages us to propose a solution for the betterment of Internet users.

We propose **'SMARTCLICK'** as the solution to handle major categories of dark patterns. Almost every Internet user uses ecommerce portals for their shopping needs and so our solution 'SMARTCLICK' targets major and most used dark patterns in ecommerce portals. 'limited users choice', 'Fake urgency & Scarcity', 'Hidden cost' and 'User reviews' are the most prevalent dark patterns in the ecommerce portals and we propose to handle these in our solution.

The aim and scope of the project is to continuously monitor the user interactions and transactional behavior, consequently detecting the dark patterns and trigger the alert response to notify users.

2. Detailed Description of the Project

Dark patterns are carefully crafted interfaces to trick people into making decisions or performing actions that they otherwise would not. They undermine users' intentions and cost them time or money they never meant to spend.

SMARTCLICK is a novel approach designed to protect the user from the harmful effects of Dark Patterns. The following are the salient features of SMARTCLICK:

- A JavaScript based Browser extension with capability of real-time monitoring of user interactions on e-commerce portals.
- Leveraging the Browser extension with the intelligence of Machine Learning algorithms on the server side to detect patterns indicative of dark design practices.
- Optimize the Machine Learning model for efficiency and ensuring computational lightness on the client side.
- Assured data privacy by the browser extension while using user data for model building.

The proposed project will involve employing state-of-the-art machine learning algorithms to build a robust Machine Learning models to predict Dark Patterns and warn the users for the same. The overall development process will consist of the following steps:

a) **Data collection:** Gather relevant data from reliable sources, including past observations. The database contains the category and the features of the dark patterns that include the approaches used to disguise the users.
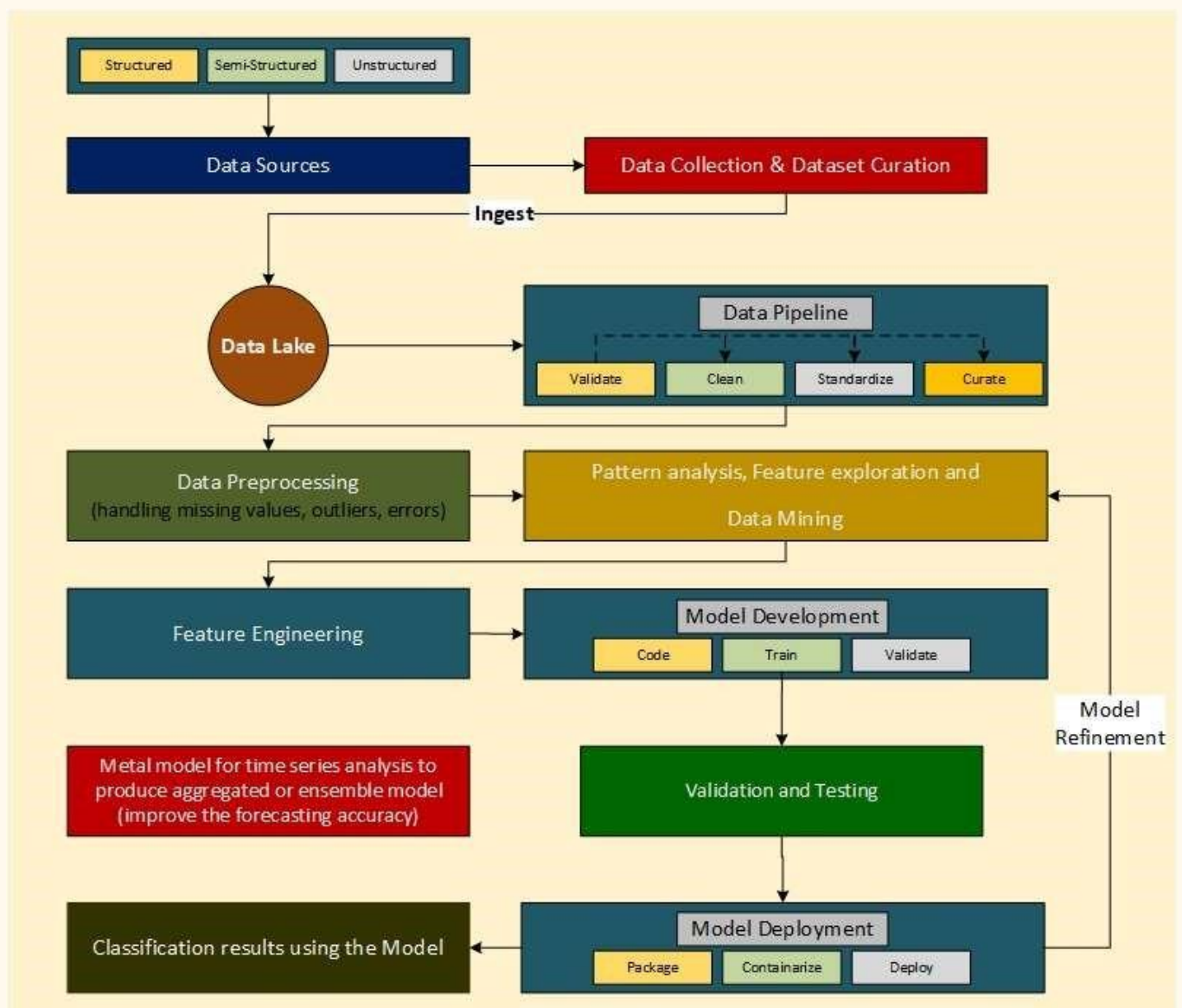

Figure 1: Block diagram representation of the holistic approach

b) **Data Preprocessing:** Cleanse the collected data, handle missing values, eliminate outliers, and perform necessary feature engineering to improve the model's performance.

c) **Pattern analysis and mining:** Before model selection the data patterns need to be analyzed by using the following methods: Pattern mining, feature analysis and then analyze the similar patterns in collected data.

d) **Model Selection:** Evaluate various Machine Learning algorithms and select the most suitable approach based on the dataset's characteristics and prediction requirements.

e) **Model Training:** Train the chosen model using the preprocessed data. Implement techniques like cross-validation, regularization, and hyper-parameter tuning to ensure optimal performance.

f) **Data prediction:** The trained AI-based model will be used to detect dark pattern in new data inputs.

g) **Model Evaluation:** Assess the model's performance by measuring prediction accuracy metrics like mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE).

h) **Model Deployment:** Develop a user-friendly interface or API for seamless integration of the prediction model into existing systems or applications.

3. An Ecosystem for Dark Patterns

History

Dark patterns have gained attention in recent years due to their prevalence in digital interfaces and their potential negative impact on users. Looking into its history, the term "dark patterns" was coined and gained prominence in the mid-2010s, though this manipulation have been around for much longer. Later the designers and researchers started to recognize and categorize these deceptive techniques as distinct and problematic aspects of user interfaces.

Prevalence

The prevalence of dark patterns expanded with the digital landscape expansion. In current era of online shopping, companies use them to influence user behavior for business gains, such as increasing sales, acquiring user data, or promoting certain features. The trending common instances are:

- **Misdirection:** Steering users towards unintended actions or choices.
- **Hidden Costs:** Concealing additional charges until later stages of a transaction.

- **Forced Continuity:** Making it difficult for users to cancel subscriptions or opt-out of services.
- **Roach Motel:** Easy to get into a situation but difficult to get out of it, like compeling a user for subscribing to a service
- **Trick Questions:** Presenting questions or options designed to lead users into unintended choices.

Significance

The significance of addressing dark patterns extends across several dimensions:

- User expectation: Dark patterns are destroying users' trust by intentional deceitful tactics, leading to negative user experiences leading harm to brand reputation.

- Ethical Concerns: Dark patterns may violate principles of transparency and user autonomy which is questionable to the use of credible techniques in interfaces. Designers and developers are prioritizing ethical design, focusing on user well-being and transparent communication.

- Social Awareness: Awareness among users has grown, resulting in addressing dark patterns through privacy and consumer protection regulations.

4. Aims and Goals

Aim

To curate a dataset of data patterns and perform data mining, pattern matching and feature exploration on the datasets to learn and build a robust classifier model to detect and warn users for data patterns.

Goals

- **Dataset Curation:** Dataset creation will help to explore and understand Dark Patterns and analyzing the various types and their characteristics, and how they operate by extracting their features. Goal is to come up with curated datasets for the common and frequent ones found in ecommerce shopping platform like 'limited users choice', 'Fake urgency & Scarcity','Hidden cost' and 'user reviews'.

- **Building a Machine Learning model:** Machine Learning focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. Machine Learning is equipped with Automated learning and Adaptability. It generalizes the patterns and extracts features to predict with accuracy. The Machine Learning model can detect deceptive patterns and execute alert mechanism on the use Browser Extension with support from the model residing at the server.

- **Awareness and Education:** Developing a scheme to raise awareness among designers, developers, and users about the existence and consequences of dark patterns by its usage.

Figure 2: Goals of the proposed project

5. Major Objectives

- SMARTCLICK is developed to empower consumers by seamlessly integrating advanced machine learning (ML) algorithms into its real-time protection mechanisms operating in the background to minimise the impact of dark patterns on the user experience.

- Implementation of SMARTCLICK, with advanced technical tools will enable real-time analysis, adaptive learning, and proactive alerting. This comprehensive approach will enhance awareness the user to stand against the misleading intentions.

6. Implementation

a) **Browser Extension:** Develop a browser extension using JavaScript to continually monitor user interactions, capturing data on page elements, user clicks, and navigation patterns on a real-time basis and perform basic analysis for data patterns detection on the client browser itself.
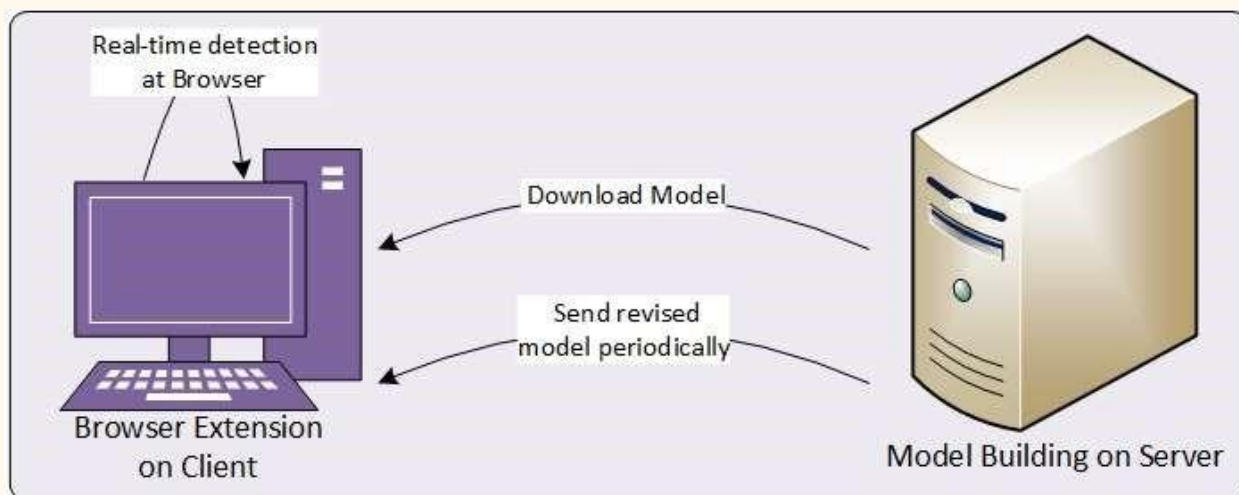

Figure 3: Architectural representation of the solution

b) **Server-Side Model Building:** The application of machine learning algorithms on the datasets which are computationally heavy tasks will be processes need to be executed on a server and the model built needs to be compressed and sent to the browser extension. The task would involve implementing machine learning algorithms (e.g., decision trees, random forests) on the server side to analyze user interactions data and train the model to recognize dark patterns.

c) **Model Integration with Browser Extension:** The model built on the server is to be periodically sent to the browser on the client side, ensuring minimal impact on browser performance.

d) **Real-Time Detection:** Enable real-time detection of dark patterns by the JavaScript based clientside component, providing immediate feedback to users when potentially manipulative design elements are encountered.
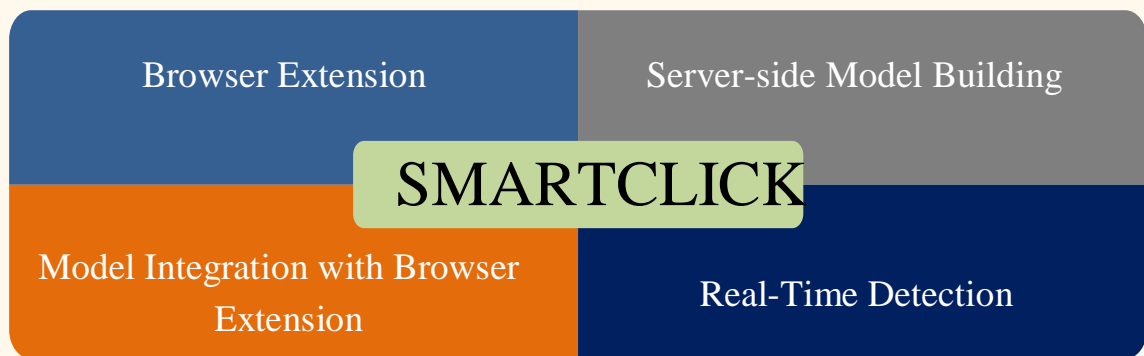
Browser Extension	Server-side Model Building
SMARTCLICK	
Model Integration with Browser Extension	Real-Time Detection

Figure 4: Feature Components of the solution

7. Need for Machine Learning

Detecting dark patterns is crucial for protecting user rights, privacy, and ensuring a fair and transparent online environment. Machine learning can play a significant role in detecting dark patterns due to its ability to analyze large datasets and equip with salient features.

- Machine Learning systems have the ability of automated learning to improve their performance over time without being explicitly programmed.

- Machine Learning models can adapt to changing environments and new data.

- Machine Learning models can generalize patterns from the training data to make predictions on new, unseen data.

- Machine Learning models can automatically extract relevant features or characteristics from raw data.

- Machine Learning algorithms can scale to handle large and complex datasets. Its flexibility is valuable in addressing real-world problems that may exhibit non-linear patterns.

- Machine Learning models provide probabilistic outputs, with their predictions to understanding the reliability of the model's decisions.

8. Development, Simulation and Benchmarking

SMARTCLICK is a holistic solution aiming to address dark patterns. Incorporation of Machine Learning in addressing patterns is imperative due to various factors aligning with the goals of this innovative solution enforcing the following developments:

Pattern Recognition and Image Analysis: Machine learning models can be trained on datasets of images and texts of known dark patterns. These algorithms can then identify similar patterns in new interfaces, helping to automatically detect potential instances of dark patterns designed to mislead users.

User Feedback and Behavior Analysis: Machine Learning models can be used to identify anomalies in user behavior. User feedback, reviews, and complaints to identify common patterns or keywords and deviations from normal user behavior, such as excessive clicking or rapid decision-making, can be indicative of dark patterns.

Natural Language Processing (NLP): Many dark patterns involve misleading or confusing language to manipulate users. NLP models can analyze the text content on websites and applications to identify deceptive or ambiguous language, helping to flag potential dark patterns.

Supervised and Continuous Learning: By utilizing labeled datasets where instances of dark patterns are identified, supervised learning algorithms can learn to classify websites or applications as either having or not having dark patterns. This can be a proactive approach to identifying potential issues. Also that the machine learning models can be continuously updated and retrained to adapt to evolving threats in the digital landscape.
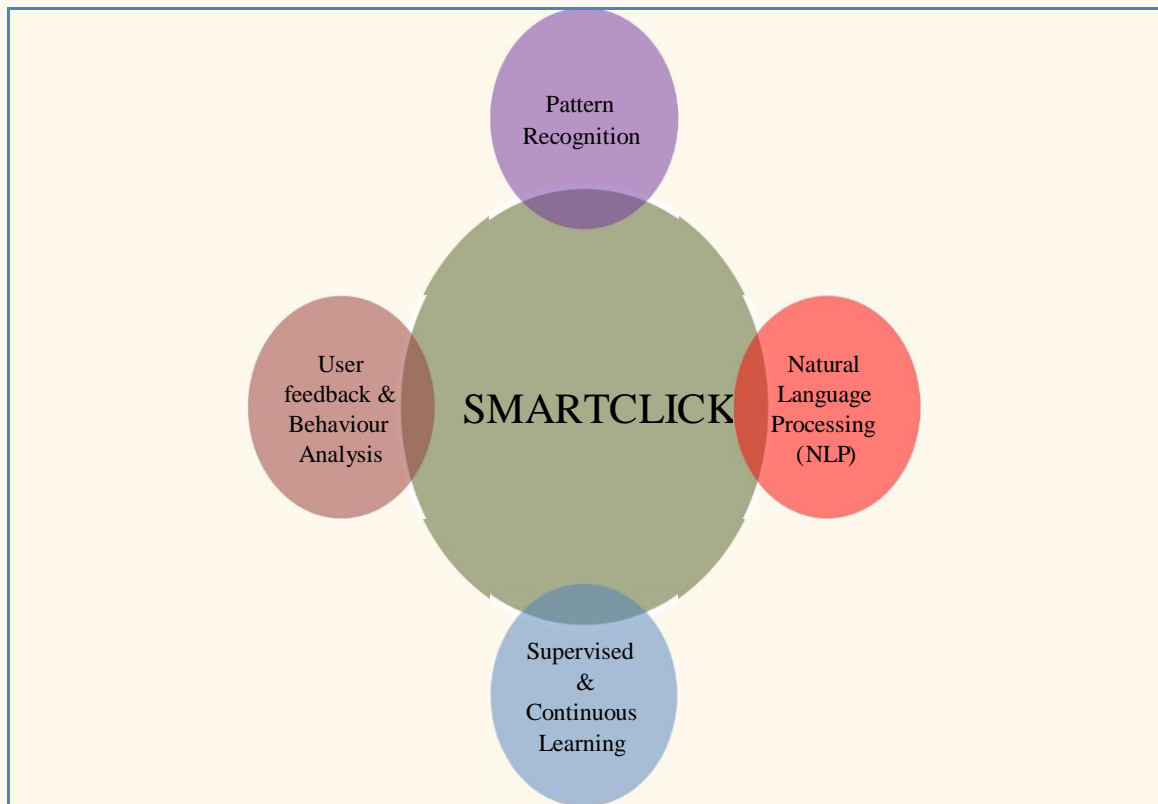
Figure 5: Developmental approaches

Benchmarking of SMARTCLICK:

- An accuracy range of 90 95% to ensure precise detection of dark patterns.

- A precision range between 85 – 90% to minimize false positives, ensuring that identified dark patterns are likely to be genuine.

- F1 score range between 85-90% to strike a balance between precision and recall.

- Real-Time Alert Response, evaluates the speed of Smart Click's alert system in providing timely notifications to users upon detecting dark pattern, enhancing user awareness and decision making.

- Cross-Browser compatibility ensures that SmartClick delivers reliable dark pattern detection irrespective of the browser used by the end user.
- Scalability enables the system to handle increased loads to aim for optimal responsiveness.

9. Generative Artificial Intelligence

SMARTCLICK an innovative solution designed to counteract dark patterns on e-commerce platforms, employs Generative Artificial Intelligence (AI) to enhance its detection capabilities. This technology enables SmartClick to recognize and generate patterns associated with deceptive design practices, ensuring the identification of variations and evolutions of known dark patterns.

- The adaptive learning feature of Generative AI allows SMARTCLICK to continuously evolve based on new data, proactively adapting to changes in the digital landscape and staying ahead of emerging deceptive tactics.

- The image and visual analysis component of Generative AI in SMARTCLICK focuses on scrutinizing visual elements on web pages, enhancing detection capabilities by recognizing visual clues indicative of dark patterns.

- User behavior analysis is strengthened by Generative AI, allowing SMARTCLICK to identify deviations aligning with known dark patterns in real-time.

- Combining supervised learning for known patterns and unsupervised learning for anomaly detection, SMARTCLICK employs Generative AI in a dual approach, enhancing its ability to identify both familiar and emerging deceptive design practices.

10. Deliverables

a) Dataset for Dark patterns
b) Features list for identifying common dark patterns
c) Classifier Model for detection of dark patterns
d) Browser extension for near real-time warning on dark patterns
e) Deployed Classifier Model

11. References

1) https://medium.com/@varun.sja/learn-how-to-build-a-machine-learning-model-in-8-steps-7e62a06526e3

2) https://analyticsindiamag.com/the-7-key-steps-to-build-your-machine-learning-model/

3) https://towardsdatascience.com/the-7-steps-of-machine-learning-2877d7e5548e

4) https://link.springer.com/article/10.1007/s42979-021-00592-x

5) https://mobidev.biz/blog/essential-machine-learning-algorithms-for-business-applications

6) https://towardsdatascience.com/machine-learning-model-deployment-as-rest-api-in-four-easy-steps-e65764f7c5c

7) https://towardsdatascience.com/implementing-ml-systems-tutorial-server-side-or-client-side-models-3127960f9244